Social Psychological and Personality Science OnlineFirst, published on August 6, 2010 as doi:10.1177/1948550610379056

Social Psychological and
Personality Science
000(00) 1-8

DOI: 10.1177/1948550610379056
http://spps.sagepub.com
SAGE

The Expression of Personality in Virtual Worlds

Nick Yee[1], Helen Harris[1], Maria Jabon[1], and Jeremy N. Bailenson[1]

Abstract

To examine the expression of personality in virtual worlds (VWs), the authors tracked the behavioral and linguistic output of 76 students continuously over a 6-week period in the VW *Second Life* (SL). Behavioral metrics in SL were consistent over time, but low stabilities were observed for linguistic metrics. To examine the ways in which personality manifested in SL, participant's Big Five scores were correlated with their virtual behavioral and linguistic metrics. For example, Conscientiousness was correlated with many metrics related to geographical movement; however, there was low overlap with findings from previous studies. The authors provide some reasons for this low concordance. Their study hints at the potential of leveraging VWs to understand the link not only between personality and behavior but also among other social and psychological phenomena as well.

Keywords

personality, quantitative models, research methods, identity, virtual worlds

Every morning, at exactly 7 o'clock, Stella treks to her farm to harvest and plant a new crop of peas, but in the popular Facebook harvesting game *FarmVille*, she can do this without even breaking a sweat. And over in the virtual world (VW) *Second Life* (SL) where users create all of the world's content, Marcus is wondering whether his new Mohawk hairdo would send the wrong message at the academic panel he is attending or whether it would be considered stylish in the context of SL. As VWs become mainstream, a critical psychological issue is whether and how personality manifests itself in VWs.

The Expression of Personality

Research in person perception has consistently shown that judgments of personality at zero acquaintance hold some degree of validity. This has been shown to be true for face-to-face encounters (Kenny, Horner, Kashy, & Chu, 1992) as well as judgments based solely on observations of an individual's bedroom or office (Gosling, Ko, Mannarelli, & Morris, 2002) or their music preferences (Rentfrow & Gosling, 2006).

Similar research has also extended to computer-mediated communication. In particular, past findings have shown that somewhat accurate personality impressions can be formed based on an individual's personal website (Marcus, Machilek, & Schutz, 2006; Vazire & Gosling, 2004), Facebook profile (Back et al., 2010), email content (Gill, Oberlander, & Austin, 2006), and even email address (Back, Schukle, & Egloff, 2008).

In exploring different methods of studying personality manifestation, some researchers (Mehl, Gosling, & Pennebaker, 2006) have illustrated the value of using natural observations to study personality as it manifests in everyday life. Collecting observations of natural behavior, however, is a daunting task where the tedium of continuous observations has only recently been offset by modern technology.

Unobtrusive Observations in Virtual Environments

VWs provide unique affordances for studying the link between personality and behavior. For the purposes of this article, we define VWs as graphical environments that enable geographically distant individuals to interact via graphical avatars (i.e., digital representations of users). These environments are no longer academic prototypes but have become mainstream interaction platforms. For example, the online game *World of Warcraft* has more than 11 million active subscribers worldwide (White, 2008). SL is another example of a VW and is unique in that users in SL create almost all the content (e.g., buildings, cars, dresses, hair styles, dance animations) in the world using scripting and modeling tools. This is in contrast to most online games where players can only use and interact with the content created by game developers.

There are three unique affordances of VWs with regard to natural observations of behavior. First, VWs are already

[1] Stanford University, Stanford, CA, USA

Corresponding Author:
Nick Yee, Stanford University, 450 Serra Mall, Stanford, CA 94305
Email: nyee@stanford.edu

instrumented with high-precision sensors. The computer systems running VWs already track the movement and behavior of every avatar to make interaction possible. Second, these high-precision sensors can track behaviors continuously and longitudinally. And finally, all these observations can be performed unobtrusively, thereby eliminating the observer effect (Webb, Campbell, Schwartz, & Sechrest, 1966).

Stability of Behavioral and Linguistic Metrics in VWs

Given that almost all studies of person perception in online environments have employed static or snapshot observations, such as personal websites (Marcus et al., 2006), the stability of behavior and linguistic output in digital environments and its relation to personality manifestation are understudied. Important clues can be inferred from other studies however. For example, written language use in diaries, class assignments, and professional journal abstracts has been shown to be quite reliable over weeks and even years (Pennebaker & King, 1999). And in a study of everyday conversational language collected via automated voice recorders (Mehl & Pennebaker, 2003) across 2-day periods separated by 4 weeks, linguistic measures were also observed to be highly consistent. Thus, these two studies show that both spoken and written language appears to be quite stable over time in physical contexts.

In one of the few longitudinal studies of behavior in virtual environments, it was found that behavioral changes over time do occur as users acclimate to interacting via digital avatars but that these changes occurred across all users (Bailenson & Yee, 2006); however, personality factors were not explored in that study. If we assume that this pattern generalizes to most VWs, then it implies that rank-order stability may be high even though absolute stability is low—that is, over time, VW users may all explore less, but some users will always tend to explore more than other users. Consequently, we seek to examine the stabilities of behavioral and linguistic output in VWs.

Research Question 1a: Are behavioral and linguistic measures in VWs stable over time in absolute terms?
Research Question 1b: Are behavioral and linguistic measures in VWs stable over time in relative terms?

Manifestation of Personality in VWs

Research in person perception has documented the ways personality manifests itself in a wide variety of environments. Here, we first consider past studies in behavioral correlates and then linguistic correlates of personality. In an early study of how personality manifested in normal face-to-face conversations (Funder & Sneed, 1993), coders rated participants in a social interaction using 64 behavioral categories. Acquaintances of participants then rated each participant using a Big Five personality inventory. Many significant correlations were observed between the personality ratings and the coded behaviors, most of which aligned closely with trait definitions of the personality factors. For example, Extraverted individuals spoke louder, with more enthusiasm and energy, and were more expressive. Agreeable individuals expressed sympathy, seemed to enjoy the interaction with their partner, and expressed interest in what their partner said.

In another study, researchers explored the manifestation of personality in personal spaces (Gosling et al., 2002). Thus, instead of observing behaviors directly, personal spaces hold identity claims (e.g., a poster of Nietzsche) and behavioral residues (e.g., a withered house plant) that reflect personality more indirectly. Nevertheless, researchers found significant correlations between coded personal space attributes and self-report personality ratings of individuals. In their study of bedrooms, it was found, for example, that individuals who scored high on Openness to Experience had more varied books and magazines. As another example, Conscientious individuals had more well-lit, neat, and well-organized bedrooms.

These two studies were selected to illustrate how direct and indirect behavioral correlates of personality have been observed in the past. It is unclear, however, how these might translate into VWs. For example, although individuals are able to interact in VWs, many VWs do not have user-controlled facial or hand gestures. And it is unclear how varied book collections translate into VWs where people do not read virtual books. On the other hand, as we mentioned above, there are a plethora of behavioral metrics that VWs provide, such as geographical movement, that may nevertheless be significant personality cues.

Unlike behavioral correlates, linguistic correlates of personality may translate more directly into VWs. To provide an overview of findings in this area, we describe four studies that have all used the Linguistic Inquiry and Word Count (LIWC) software (Pennebaker, Booth, & Francis, 2007) to examine linguistic correlates of personality. LIWC is a dictionary-based word count tool that counts the ratio of words in 70 linguistic categories. For example, the positive emotion category contains the words *happy*, *cheerful*, *joy*, and so on. The four selected studies span the past decade and examine different linguistic content: personal writing profiles (Pennebaker & King, 1999), self-narratives (Hirsh & Peterson, 2009), everyday conversations (Mehl et al., 2006), and blog content (Yarkoni, in press).

To provide a concise and coherent summary of the findings (oftentimes with hundreds of comparisons) without being bogged down by idiosyncratic differences, we summarize only correlates that were found to be significant in at least two studies (see Table 1). Of note, most linguistic correlates mirror trait definitions. For example, Emotional Stability is negatively correlated with negative emotions and Agreeableness is positively correlated with social involvement and positive emotions. On the other hand, grammatical features such as articles and first-person singulars also made frequent appearances (both were significant correlates in three out of the five personality factors), although their connection with personality is less obvious.

Thus, in the present study, we were interested in the following research question:

Research Question 2: What behavioral and linguistic correlates of personality in VWs?

Table 1. A Summary of Previous Linguistic Correlates With Personality Factors

Personality factor	Positive correlates	Negative correlates
Emotional Stability	Articles	Anger, anxiety, negative emotions, first-person singular
Extraversion	Social processes, positive emotions	
Openness to Experience	Articles, exclusives	First-person singular, present tense, past tense, social processes
Agreeableness	First-person singular, inclusives, family, positive emotions	Articles, anger, negative emotions
Conscientiousness	Achievement	Causation, exclusives, anger, negations, negative emotions

Method

Participants

As part of a class that was largely designed around their participation in SL, 76 undergraduate and master's students (25 female) participated in the study. None of the participants had previous experience with SL. Over the course of the 6-week study, the average number of hours spent in SL was 36.03 ($SD = 5.27$). The mean age of the participants was 21.07 ($SD = 3.68$).

Procedures

Prior to the start of the experiment, all participants were required to attend a 1-hour tutorial in which they were taught the basics of SL. On creating their new avatar, they visited the experimenters in SL to receive two items. First, each participant was given 1,000 Linden dollars (L$1000) to use as they wanted. Then, a *Sender* object was transferred to the participant, and the experimenter confirmed that it was attached to the avatar. The *Sender* was developed using SL's scripting language. When attached to the avatar, the tool gathered data on movement, action, and chat every 10 seconds and then transmitted the information to a database. The details of the script used to create the *Sender* are described in previous work (Yee & Bailenson, 2008). Participants were asked to spend at least 6 hours each week in SL, and their behaviors in SL were logged for 6 weeks. Participants did not have access to the database where their logged data were stored.

Measures

Personality measures. A 50-item scale measuring the Big Five factor structure was drawn from the International Personality Item Pool (Goldberg, 1999). Participants rated themselves on the inventory items using a scale that ranged from 1 (*very inaccurate*) to 5 (*very accurate*). The alpha reliabilities for Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness to Experience were .90, .70, .72, .85, and .78, respectively.

Virtual behavior metrics. The *Sender* collected longitudinal data of participants' behaviors over the 6-week period. These behaviors included their current stance (e.g., walking, flying, sitting), their Cartesian coordinate in the world, the number of other avatars within a 20 m radius, how often they logged on SL, and whether they were typing. The world of SL is divided into many different square zones with touching edges. SL users are thus always located in one particular zone, and this was also recorded by the *Sender*. Of note, it is also possible for SL users to fly as a form of locomotion. From these raw data points, we generated 17 behavioral metrics and calculated the means for these metrics for each week. In Table 2, we show the averaged weekly metrics over the 6 weeks.

Linguistic measures. The *Sender* also collected the frequency that participants used the in-world text chat and the content of their text messages. These text messages were analyzed using the word-based language analysis program LIWC (Pennebaker et al., 2007). LIWC outputs the ratio of total words that fall into each of the 70 categories (e.g., first-person singular pronouns). Past research in linguistic analysis of personality projection in everyday life has identified a set of 23 variables of interest. For the current study, we used these variables for the analysis, except for two variables associated with spoken language but not typed chat (i.e., nonfluencies such as *ummm* and filler words such as *yaknow*, *Imean*). We also included "future tense," as both "present tense" and "past tense" were included as variables of interest in the past. Finally, we added one chat frequency variable collected by the *Sender*—a count of the number of chat lines (which is different from the count of all words). Thus, we included 23 linguistic variables in the analysis. In Table 3, we show the descriptions and the averaged weekly means and standard deviations for these linguistic variables over the 6 weeks of the study. Note that LIWC outputs ratios from 0 to 100 (i.e., 50 is equal to 50%), and we follow this format in showing the means in the table.

Results

Stability of Behavioral and Linguistic Metrics in VWs Over Time

To examine the absolute stability of the VW metrics over time, we conducted repeated measures ANOVAs for each measure over the six-week period. The resulting F values and corresponding significance levels are shown in Tables 2 and 3. Among the behavioral metrics, 16 of the 17 metrics were significantly different over time. Among the linguistic metrics, 7 of the 23 were significantly different over time.

Table 2. Description of the Behavioral Metrics, Along With Their Respective Means, Standard Deviations, Averaged Pairwise Correlations Over the 6 Weeks, and the *F* Value From the Repeated Measures ANOVA Over the 6 Weeks

Variable	Description	*M*	*SD*	*r*	*F*
Fly	Ratio of time flying	0.07	0.11	.40*	5.38*
Walk	Ratio of time walking	0.06	0.06	.26*	9.45*
Run	Ratio of time running	0.03	0.08	.33*	5.84*
Sit	Ratio of time sitting	0.40	0.30	.49*	17.02*
Type	Ratio of time typing	0.02	0.02	.30*	4.21*
Teleports	Number of teleports	15.87	12.66	.37*	30.91*
Favorite zone	Ratio of time spent in the participant's most visited zone	0.55	0.28	.32*	29.74*
Unique zones	Number of unique zones visited	15.03	13.59	.41*	26.76*
Log-ins	Number of times the participant logged in	11.51	8.75	.26*	11.78*
Log-in time	Average time the participant spent in SL each unique log-in (in minutes)	49.11	56.11	.30*	10.77*
Total in radius	Average number of other avatars in the participant's 20 m radius	4.19	3.77	.41*	3.10*
Max in radius	The maximum number of other avatars in the participant's 20 m radius	13.46	4.62	.22	3.83*
Zero in radius	Ratio of total time with no other avatar within a 20 m radius	0.35	0.29	.36*	1.85
Total distance	Total distance traveled (in SL meters)	6224.71	6639.80	.39*	15.82*
Walked distance	Distance walked (in SL meters)	3777.53	4522.50	.34*	6.46*
Flown distance	Distance flown (in SL meters)	2447.18	3652.28	.42*	14.49*
Zone crossings	Number of zone crossings made	23.10	22.28	.41*	20.73*

Note: SL = *Second Life*.
* $p < .05$.

Table 3. Description of the Linguistic Metrics, Along With Their Respective Means, Standard Deviations, Averaged Pairwise Correlations Over the 6 Weeks, and the *F* Value From the Repeated Measures ANOVA Over the 6 Weeks

Variable	Description	*M*	*SD*	*r*	*F*
Chat lines	Number of chat lines	65.50	62.70	.34*	0.56
Word count	A count of all the words in the text messages sent	260.12	215.84	.25*	7.13*
Words w/ more than 6 letters	A count of all words with more than 6 letters	7.18	3.99	.15	0.24
First-person singular pronouns	I, me, my	5.15	2.19	.12	3.53*
First-person plural pronouns	We, us, our	0.36	0.33	.07	1.82
Total second-person pronouns	You, your	3.06	2.23	.32*	0.99
Total third-person pronouns	She, him, their	0.48	0.44	.06	1.12
Negations	No, not, never	1.52	0.85	.05	0.64
Articles	A, an, the	2.71	1.35	.15	2.06
Prepositions	To, with, above	5.95	2.60	.24*	2.93*
Swear words	Damn, bastard	0.55	0.85	.20	0.56
Positive emotions	Happy, good	6.54	3.23	.13	0.88
Negative emotions	Hate, ugly	1.66	1.58	.12	0.86
Causation	Because, effect	1.69	1.09	.04	1.35
Insight	Realize, know	1.27	0.75	.14	3.26*
Discrepancy	Would, should	1.04	0.74	.08	1.06
Tentative	Perhaps, maybe	1.92	1.18	.13	3.35*
Social processes	Friend, talk	11.86	6.28	.05	1.73
Past tense	Was, went	1.80	1.15	.17	0.60
Present tense	Is, go	11.71	4.44	.16	6.95*
Future tense	Will be, will go	0.56	0.46	.13	1.63
Inclusive	With, and	1.49	0.76	.07	2.68*
Exclusive	Except, but	1.57	0.85	.11	1.55

Note: Means are weekly averages. Thus, the overall 6-week average number of chat lines produced is 6 times what is listed as the weekly mean.
* $p < .05$.

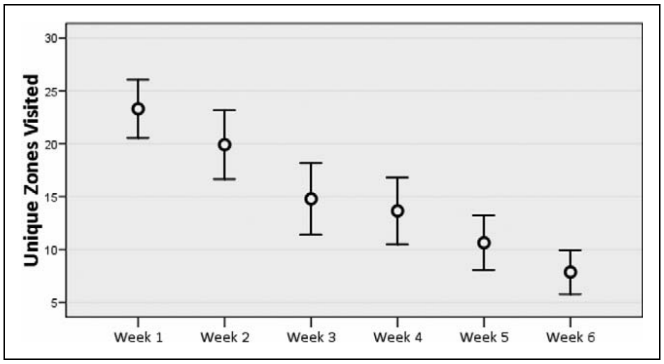

Figure 1. Average unique zones visited by week with 95% confidence interval error bars

To examine the rank-order stability of the VW metrics over time, we calculated the average of all pairwise correlations for each variable over the 6 weeks. The resulting average correlation coefficient is also shown in Table 2 and 3. Among the behavior metrics, 16 of the 17 correlation coefficients were significant, and the overall average correlation r was .35. Among the linguistic metrics, 4 of the 23 correlation coefficients were significant, with an overall average correlation r of .14.

This suggests that behavioral metrics in SL have high rank-order stability even though the absolute stability is low. As an illustration of these changes over time, the plot of visits to unique zones (i.e., the number of different areas of the SL world that a user walked through) over the 6-week period shows a significant linear contrast, $F(1, 74) = 117.48$, $p < .001$ (see Figure 1). Although the repeated measures ANOVAs suggest that virtual behavioral metrics change over time, the averaged pairwise correlations suggest that they nevertheless capture relative individual differences beneath these global trends.

On the other hand, the analyses show that linguistic content in SL is not very stable in absolute or relative terms. Nevertheless, given that we collected data over a 6-week period, there is still the potential that the mild stabilities in linguistic output aggregated over time may yield some markers of personality expression.

Markers of Personality in VWs

Given the increased risk of experiment-wise error in large correlation tables with 39 variables against the Big Five factors, we used an analytic method developed by Sherman and Funder (2009) to address this specific issue. The method employs a Monte Carlo simulation of repeatedly randomized data within each participant. Thus, the method preserves the statistical properties of the data gathered. The method conducts 1,000 of these randomized data sets and tabulates the number of observed significant correlations (at α of .05). The probability of the actual number of significant correlations is then calculated based on where it lies on the distribution of the 1,000 randomizations. In our case, using an alpha of .05, we had 26 observed significant correlations, where only 10.68 would be expected by chance based on the simulations. According to this Monte Carlo method, the probability of this number of observed correlations is $p = .02$. This provides assurance that the observed correlations, as a set, are nonrandom. We present the significant correlations in Table 4 and compare them to our summary of previous findings mentioned in the introduction.

Here we briefly highlight some clusters of related correlates in Table 4. Conscientiousness was correlated with many variables related to geographical movement, such as distance walked and number of unique zones visited. Emotional Stability was related to log-in patterns—participants low on Emotional Stability logged in more often but with shorter durations. Participants low on Emotional Stability were also more likely to reduce their use of common linguistic features, such as pronouns and tense markers. On the other hand, it is more difficult to make sense of the correlates for the other three personality factors; there was only one significant correlate for Openness and two correlates for Agreeableness. The correlates for Extraversion are also difficult to interpret as a whole.

Discussion

Our study illustrates that the affordances of VWs can be leveraged to collect large amounts of detailed longitudinal behavioral and linguistic metrics from these environments unobtrusively. Overall, although our findings suggest that personality is expressed in VWs via both behavioral and linguistic correlates, the specific correlations we found did not match those identified in previous studies, nor were they easily interpretable for the most part.

Similar to the pattern observed in a longitudinal study of interactions in a virtual environment (Bailenson & Yee, 2006), we found that although absolute stability was low in terms of behavioral metrics in VWs, the rank-order stability was robust. On the other hand, the stability of linguistic metrics in VWs was low. This is in contrast to a study of linguistic stability in everyday life (Mehl & Pennebaker, 2003). One potential explanation is that the varied settings (e.g., teleporting from a poetry reading to a disco party) in SL introduced a high amount of noise into the linguistic metrics. In contrast, linguistic interactions in everyday life are often part of a routine and familiar to the individual. Although it is not clear what impacted the stability of linguistic metrics in our study, some of the aggregate linguistic measures over the 6-week period did correlate well with personality factors, so it appears that markers of personality expression can be derived from linguistic metrics in VWs by aggregating over a large period of time.

Significant behavioral and linguistic correlates were found for all the Big Five factors; however, our findings do not match up well with findings from the previous studies summarized in the introduction. In comparing our findings to the summarized findings in Table 1, we find only one match between the current data set and previous findings, the correlation between common linguistic features and emotional stability. It is troubling that the correlations from the current study did not replicate those found in much of the previous work. On the other hand, none of the findings in this study directly contradict (i.e., show a significant opposite signed correlation) those in Table 1.

Table 4. Significant Correlations Between Personality Factors and the Behavioral and Linguistic Metrics

	Big Five A	Big Five C	Big Five E	Big Five ES	Big Five O
Behavioral measures					
Walk	.19	.33*	.13	.04	.23*
Teleports	.18	.29*	.28*	–.11	.08
Favorite zone	–.07	–.26*	–.15	.10	–.18
Unique zones	.17	.26*	.17	–.02	–.01
Log-ins	.07	.15	.07	–.23*	.00
Log-in time	.08	–.05	–.16	.28*	.10
Zero in radius	.13	.26*	.09	.05	.15
Total distance	.25*	.16	.06	–.04	.00
Walked distance	.24*	.23*	.11	–.02	.03
Zone crossings	.19	.24*	.14	–.03	–.03
Linguistic measures					
Words w/ more than 6 letters	.15	.34*	.24*	.11	.05
First-person singular	.08	–.07	.22*	–.23*	.04
Total second-person pronouns	.04	.21	–.06	–.31*	–.10
Swear words	–.08	–.11	–.29*	–.06	–.09
Causation	.04	.03	.00	–.24*	–.01
Discrepancy	–.09	–.09	–.14	–.25*	–.14
Tentative	–.14	.25*	.18	–.07	–.08
Present tense	–.12	.09	.13	–.24*	.02
Future tense	.19	.04	.09	–.30*	.02
Inclusive	–.03	.04	.12	–.26*	.05
Exclusive	–.06	–.01	.26*	–.06	–.05

* $p < .05$.

After observing the low concordance between our linguistic correlates and those observed in previous studies, we skimmed through the actual logged chats and noticed several unique features of chat in SL. Our first clue that chat on SL was very different from the comparison linguistic samples in previous studies was the observation from Table 3 that the average chat line contained only about four words. Perusal of the actual chat logs also revealed that SL users often employed abbreviations such as *U* for *you* and *rly* for *really*. They also employed many common internet acronyms such as *lol* for *laughing out loud* and emoticons such as *<3* for the heart symbol. We also observed frequent typos, sentence fragments, and pronoun drops (e.g., *busy now* instead of *I'm busy now*). These unique linguistic features of chat in SL likely contributed to both the low stability of the linguistic measures as well as their nonconcordance with findings from previous studies. Indeed, scholars have noted that instant messaging has its own lexicon, grammar, and usage conditions and is distinct from written prose and normal speech (Crystal, 2001; Ling & Baron, 2007; Tagliamonte & Denis, 2008; Walther, Gay, & Hancock, 2005). Our findings highlight the fact that although many Web 2.0 systems provide a wealth of linguistic data, it is important for researchers to develop linguistic tools that can take into account the unique linguistic aspects of communication in these novel environments.

Despite the limitation of the chat data, the current data are exciting because they add the use of behavioral and nonverbal data as a tool to examine personality. As Table 4 demonstrates, a number of features based on locomotion and geography are significant correlates of personality, especially conscientiousness. Although the linguistic data may be limited based on the specific chat setup in SL, the richness of the behavioral data provides unique insights.

There were several other limitations to the study. First and foremost, our study focused on only one VW. Although some personality cues may appear in other VWs, it is at present not clear how many of our findings generalize to other VWs. Second, only undergraduate students were included in the study sample, and this too may limit its generalizability. And finally, in hindsight, the behavioral metrics we logged largely centered on variations of geographical movement and in turn may have constrained the manifestation of personality in VWs we could identify. For example, we tracked movement through different zones but had no good way of coding for the content of those zones or the context of social gatherings. The affordances of different VWs or creating tools to extract more contextual data may reveal other behavioral correlates.

Nevertheless, our findings do show that there are significant manifestations of personality in VWs. The behavioral correlates suggest that Conscientiousness is related to geographical movement in VWs and that Emotional Stability is related to log-in patterns. Overall, we believe that the current study presents a first step in understanding personality expression in the novel domain of VWs. As Mehl and his colleagues (2006) noted, capturing people's interactions in the physical world reveals what people spontaneously do, what they avoid, and their idiosyncrasies. The value in natural observations lies in its ability to "document personality right where it occurs" (p. 875). Although natural observations of virtual behavior may seem ironic at first glance, it is important to remember that much of our daily lives now takes place in virtual places. For example, the average online gamer spends more than 20 hours a week in their game avatars (Williams, Yee, & Caplan, 2008; Yee, 2006). More importantly, the viability of longitudinal behavioral tracking leveraging VWs as a methodology extends well beyond the domain of personality psychology. Indeed, it is not difficult to imagine using a similar methodology to examine the emergence or stability of social norms, leadership, or stereotypes. In sum, VWs offer both new methods and new contexts to understand the links between psychological factors and social phenomena.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

Funding

The current work was partially supported by a grant from the National Science Foundation (NSF, grant 0527377).

References

Back, M., Schukle, S., & Egloff, B. (2008). How extraverted is honey.bunny77@hotmail.de? Inferring personality from e-mail addresses. *Journal of Research in Personality*, *42*, 1116-1122.

Back, M., Stopfer, J., Vazire, S., Gaddis, S., Schmukle, S., Egloff, B., & Gosling, S. D. (2010). Facebook profiles reflect actual personality not self-idealization. *Psychological Science*, *21*, 372-374.

Bailenson, J., & Yee, N. (2006). A longitudinal study of task performance, head movements, subjective report, simulator sickness, and transformed social interaction in collaborative virtual environments. *Presence: Teleoperators and Virtual Environments*, *15*, 699-716.

Crystal, D. (2001). *Language and the Internet*. Cambridge, UK: Cambridge University Press.

Funder, D., & Sneed, C. (1993). Behavioral manifestations of personality: An ecological approach to judgmental accuracy. *Journal of Personality and Social Psychology*, *64*, 479-490.

Gill, A., Oberlander, J., & Austin, E. (2006). Rating e-mail personality at zero acquaintance. *Personality and Individual Differences*, *40*, 497-507.

Goldberg, L. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 7-28). Tilburg, Netherlands: Tilburg University Press.

Gosling, S., Ko, S., Mannarelli, T., & Morris, M. (2002). A room with a cue: Judgments of personality based on offices and bedrooms. *Journal of Personality and Social Psychology*, *82*, 379-398.

Hirsh, J., & Peterson, J. (2009). Personality and language use in self-narratives. *Journal of Research in Personality*, *43*, 524-527.

Kenny, D., Horner, C., Kashy, D., & Chu, L. (1992). Consensus at zero acquaintance: Replication, behavioral cues, and stability. *Journal of Personality and Social Psychology*, *62*, 88-97.

Ling, R., & Baron, N. (2007). Text messaging and IM: Linguistic comparison of American college data. *Journal of Language and Social Psychology*, *26*, 291-298.

Marcus, B., Machilek, F., & Schutz, A. (2006). Personality in cyberspace: Personal web sites as media for personality expressions and impressions. *Journal of Personality and Social Psychology*, *90*, 1014-1031.

Mehl, M., Gosling, S., & Pennebaker, J. (2006). Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology*, *90*, 862-877.

Mehl, M., & Pennebaker, J. (2003). The sounds of social life: A psychometric analysis of students' daily social environment and natural conversations. *Journal of Personality and Social Psychology*, *84*, 857-870.

Pennebaker, J., Booth, R., & Francis, M. (2007). Linguistic Inquiry and Word Count [Computer software]. Mahwah, NJ: Erlbaum.

Pennebaker, J., & King, L. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, *77*, 1296-1312.

Rentfrow, P., & Gosling, S. (2006). Message in a ballad: The role of music preferences in interpersonal perception. *Psychological Science*, *17*, 236-242.

Sherman, R., & Funder, D. (2009). Evaluating correlations in studies of personality and behavior: Beyond the number of significant findings to be expected by chance. *Journal of Research in Personality*, *43*, 1053-1063.

Tagliamonte, S., & Denis, D. (2008). Linguistic ruin? LOL! Instant messaging and teen language. *American Speech*, *83*, 3-34.

Vazire, S., & Gosling, S. (2004). e-perceptions: Personality impressions based on personal websites. *Journal of Personality and Social Psychology*, *87*, 123-132.

Walther, J., Gay, G., & Hancock, J. (2005). How do communication and technology researchers study the internet? *Journal of Communication*, *55*, 632-657.

Webb, E., Campbell, D., Schwartz, R., & Sechrest, L. (1966). *Unobtrusive measures: Non-reactive research in the social sciences*. Chicago, IL: Rand McNally.

White, P. (2008, May 30). *MMOGData: Charts. 2009*. Retrieved from http://mmodata.blogspot.com.

Williams, D., Yee, N., & Caplan, S. (2008). Who plays, how much, and why? Debunking the stereotypical gamer profile. *Journal of Computer-Mediated Communication*, *13*, 993-1018.

Yarkoni, T. (in press). Personality in 100,000 words: A large scale analysis of personality and word use among bloggers. *Journal of Research in Personality*.

Yee, N. (2006). The demographics, motivations, and derived experiences of users of massively multi-user online graphical environments. *Presence: Teleoperators and Virtual Environments*, *15*, 309-329.

Yee, N., & Bailenson, J. (2008). A method for longitudinal behavioral data collection in Second Life. *Presence: Teleoperators and Virtual Environments*, *17*, 594-596.

Bios

Nick Yee is a post-doctoral student in the Department of Communication at Stanford University. He studies self-representation and social interaction in virtual environments.

Helen Harris is a graduate student at Stanford University studying human-computer interaction and virtual worlds.

Maria Jabon completed her master's degree in Electrical Engineering at Stanford University and is now a software developer for LinkedIn. She is interested in machine learning and web programming.

Jeremy N. Bailenson is an associate professor in the Department of Communication at Stanford University and the founding director of the Virtual Human Interaction Lab.